

# Eléments de statistique

## Liste de questions de théorie

Louis Wehenkel

Département d'Electricité, Electronique et Informatique - Université de Liège  
B24/II.93 - L.Wehenkel@ulg.ac.be



MATH0487-1 : 3Baclng, 3Baclnf - 28/10/2015

Find slides: <http://montefiore.ulg.ac.be/~lwh/Stats/>

## Avertissement

- ▶ La liste de questions qui suit est indicative.
- ▶ Il s'agit surtout d'une aide pour structurer l'étude de la matière, et une indication quant au style des questions qui pourront être posées à l'examen écrit de théorie.
- ▶ A titre illustratif, sur la page web du cours, on pourra trouver les deux questions de théorie telles que posées lors de l'examen de janvier 2014.

NB: Outre les slides des leçons théoriques, il est vivement conseillé de lire les deux articles de référence (voir page web du cours) et d'écouter la vidéo sur le paradoxe de Simpson [https://www.youtube.com/watch?v=vs\\_Zzf\\_vL2I](https://www.youtube.com/watch?v=vs_Zzf_vL2I). Certaines sous-questions peuvent en effet porter sur cette matière.

Statistiques descriptives

Théorie de l'échantillonnage

Théorie de l'estimation

Théorie des tests d'hypothèses

# Statistiques descriptives

**SD1** Savoir donner les définitions des principales statistiques descriptives unidimensionnelles,  $m_x$ ,  $s_x$ , Médiane $_x$ , percentiles. Calculer leurs valeurs pour un échantillon donné de petite taille, et dessiner la fonction de répartition empirique  $\hat{F}_x$ .

# Théorie de l'échantillonnage

- EC1** Définir la notion d'échantillon i.i.d. d'une variable  $\mathcal{X}$  et de taille  $n$ . Discuter la distribution d'échantillonnage de la moyenne d'échantillon  $m_x$ : espérance, variance, loi en régime asymptotique, loi dans le cas où  $\mathcal{X}$  est  $\mathcal{N}(\mu; \sigma^2)$ .
- EC2** Discuter la distribution d'échantillonnage de  $f_1$  et de  $\hat{F}_x$ , en comprenant bien la notion de  $D_{KS}$ , et en énonçant de façon correcte le théorème de Glivenko-Cantelli. Dessiner  $F_{\mathcal{X}}$  pour une variable de Bernoulli de paramètre  $p_1$ , et la fonction empirique  $\hat{F}_x$  pour un échantillon  $n_0, n_1$  de cette variable; calculer  $D_{KS}(\hat{F}_x, F_{\mathcal{X}})$ .
- EC3** Discuter la distribution d'échantillonnage de  $s_x^2$ , et celle de  $s_{n-1}^2$ , d'abord en général, puis dans le cas où la variable parente est  $\mathcal{N}(\mu; \sigma^2)$ .

# Théorie de l'estimation

- ES1** Définir la notion d'estimateur ponctuel d'un paramètre  $\theta \in \mathbb{R}$  de la loi  $f_{\mathcal{X}}(\cdot; \theta)$  d'une variable aléatoire. Définir la notion de biais et de variance, et montrer que l'erreur quadratique d'estimation vaut,  $\forall \theta$ , la somme du biais au carré et de la variance. On vous demandera d'indiquer biais et variance, d'un estimateur introduit au cours.
- ES2** Expliquer le principe général de la méthode du maximum de vraisemblance, puis celui de la méthode bayésienne, et discuter leurs similarités et leurs différences. Appliquer les deux méthodes à un cas concret d'une variable de Bernoulli et discuter la différence entre les estimateurs obtenus.
- ES3** Expliquer le principe de la construction d'un intervalle de confiance, dans un des cas présentés au cours. Donner l'interprétation de la notion d'intervalle de confiance, et expliquer la différence avec les intervalles de "probabilité" de l'approche bayésienne.

# Théorie des tests d'hypothèses

- TH1** Expliquer le principe de la théorie de la décision bayésienne. Illustrez son application au problème de test d'hypothèses statistiques. Décrire le principe général de la démarche de façon formelle. Discuter les forces et les faiblesses de cette approche, sur le plan conceptuel et pratique.
- TH2** Poser le problème adressé par la méthode de Neyman-Pearson. Énoncer mathématiquement le théorème de Neyman-Pearson, **et discuter intuitivement ce qu'il nous apprend**. Définir la notion de statistique exhaustive, discuter son application aux tests d'hypothèses statistiques. Illustrer ces idées pour une situation théorique vue au cours (p.ex. loi Gaussienne ou de Bernoulli de la variable parente).
- TH3** Décrire les principaux types de test d'hypothèse vus au cours de façon synthétique (tests sur la moyenne, tests d'ajustement, tests d'indépendance). Expliquer et discuter la problématique des tests multiples.